WE CLAIM:

1.    A protein derived from an enterically transmitted non-A/non-B viral hepatitis agent whose genome contains a region which is homologous to a coding region of the 1.33 kb DNA EcoRI insert present in plasmid pTZKF1(ET1.1) carried in E. coli strain BB4 and having ATCC deposit no. 67717.

2.    The protein of claim 1, which is encoded by a complete coding region within said 1.33 kb EcoRI insert.

3.    A recombinant protein derived from an enterically transmitted nonA/nonB viral hepatitis agent whose genome contains a region which is homologous to a coding region of a DNA molecule having a first sequence (SEQ ID NO.1):

```
AGACCTGTCC CTGTTGCAGC TGTTCTACCA CCCTGCCCCG AGCTCGAACA GGGCCTTCTC    60
TACCTGCCCC AGGAGCTCAC CACCTGTGAT AGTGTCGTAA CATTTGAATT AACAGACATT   120
GTGCACTGCC GCATGGCCGC CCCGAGCCAG CGCAAGGCCG TGCTGTCCAC ACTCGTGGGC   180
CGCTACGGCG GTCGGACAAA GCTCTACAAT GCTTCCGACT CTGATGTTCG CGACTCTCTC   240
GCCCGTTTTA TCCCAGCCAT TGGCCCCGTA CAGGTTACAA CTTGTGAATT GTACGAGCTA   300
GTGGAGGCCA TGGTGGAGAA GGGCCAGGAT GGCTCCGCCG TCCTTGAGCT TGATCTTTGC   360
AACCGTGACG TGTCCAGGAT CACCTTCTTC CAGAAAGATT GTAACAAGTT CACCACAGGT   420
GAGACCATTG CCCATGGTAA AGTGGGCCAG GGCATCTCGG CCTGGAGCAA GACCTTCTGC   480
GCCCTCTTTG GCCCTTGGTT CCGCGCTATT GAGAAGGCTA TTCTGGCCCT GCTCCCTCAG   540
GGTGTGTTTT ACGGTGATGC CTTTGATGAC ACCGTCTTCT CGGCGGCTGT GGCCGCAGCA   600
AAGGCATCCA TGGTGTTTGA GAATGACTTT TCTGAGTTTG ACTCCACCCA GAATAACTTT   660
TCTCTGGGTC TAGAGTGTGC TATTATGGAG GAGTGTGGGA TGCCGCAGTG GCTCATCCGC   720
CTGTATCACC TTATAAGGTC TGCGTGGATC TTGCAGGCCC CGAAGGAGTC TCTGCGAGGG   780
TTTTGGAAGA AACACTCCGG TGAGCCCGGC ACTCTTCTAT GGAATACTGT CTGGAATATG   840
GCCGTTATTA CCCACTGTTA TGACTTCCGC GATTTTCAGG TGGCTGCCTT TAAAGGTGAT   900
```

GATTCGATAG TGCTTTGCAG TGAGTATCGT CAGAGTCCAG GAGCTGCTGT CCTGATCGCC    960

GGCTGTGGCT TGAAGTTGAA GGTAGATTTC CGCCCGATCG GTTTGTATGC AGGTGTTGTG   1020

GTGGCCCCCG GCCTTGGCGC GCTCCCTGAT GTTGTGCGCT TCGCCGGCCG GCTTACCGAG   1080

AAGAATTGGG GCCCTGGCCC TGAGCGGGCG GAGCAGCTCC GCCTCGCTGT TAGTGATTTC   1140

CTCCGCAAGC TCACGAATGT AGCTCAGATG TGTGTGGATG TTGTTTCCCG TGTTTATGGG   1200

GTTTCCCCTG GACTCGTTCA TAACCTGATT GGCATGCTAC AGGCTGTTGC TGATGGCAAG   1260

GCACATTTCA CTGAGTCAGT AAAACCAGTG CTCGA                               1295

a second sequence (SEQ ID NO.5):

TCGAGCACTG GTTTTACTGA CTCAGTGAAA TGTGCCTTGC CATCAGCAAC AGCCTGTAGC     60

ATGCCAATCA GGTTATGAAC GAGTCCAGGG GAAACCCCAT AAACACGGGA AACAACATCC    120

ACACACATCT GAGCTACATT CGTGAGCTTG CGGAGGAAAT CACTAACAGC GAGGCGGAGC    180

TGCTCCGCCC GCTCAGGGCC AGGGCGCCAA TTCTTCTCGG TAAGCCGGCC GGCGAAGCGC    240

ACAACATCAG GGAGCGCGCC AAGGCCGGGG GCCACGACAA CACCTGCATA CAAACCGATC    300

GGGCGGAAAT CTACCTTCAA CTTCAAGGCA CAGCCGGCGA TCAGGACAGC AGCTCCTGGA    360

CTCTGACGAT ACTCACTGCA AAGCACTATC GAATCATCAC CTTTAAAGGC AGCCACCTGA    420

AAATCGCGGA AGTCATAACA GTGGGTAATA ACGGCCATAT TCCAGACAGT ATTCCATAGA    480

AGAGTGCCGG GCTCACCGGA GTGTTTCTTC CAAAACCCTC GCAGAGACTC CTTCGGGGCC    540

TGCAAGATCC ACGCAGACCT TATAAGGTGA TACAGGCGGA TGAGCCACTG CGGCATCCCA    600

CACTCCTCCA TAATAGCACA CTCTAGACCC AGAGAAAAGT TATTCTGGGT GGAGTCAAAC    660

TCAGAAAAGT CATTCTCAAA CACCATGGAT GCCTTTGCTG CGGCCACAGC CGCCGAGAAG    720

ACGGTGTCAT CAAAGGCATC ACGGTAAAAC ACACCCTGAG GGAGCAGGGC CAGAATAGCC    780

TTCTCAATAG CGCGGAACCA AGGGCCAAAG AGGGGGCAGA AGGTCTTGCT CCAGGCCGAG    840

ATGCCCTGGC CCACTTTACC ATGGGCAATG GTCTCACCTG TGGTGAACTT GTTACAATCT    900

TTCTGGAAGA AGGTGATCCT GGACACGTCA CGGTTGCAAA GATCAAGCTC AAGGACGGCG    960

GAGCCATCCT GGCCCTTCTC GACCATGGCC TCCACTAGCT CGTACAATTC ACAAGTTGTA   1020

ACCTGTACGG GGCCAATGGC CGGGATAAAA CGGGCGAGAG AGTCGCGAAC ATCAGAGTGG   1080

GAAGCATTGT AGAGCTTTGT GCGACCGCCG TAGCGGCCCA CGAGTGTGGA CAGCACGGCC   1140

TTGCGCTGGC TCGGGGCGGC CATGCGGCAG TGCACAATGT CTGTTAATTC AAATGTTACG   1200

ACACTATCAC AGGTGGTGAG CTCCTGGGGC AGGTAGAGAA GGCCCTGTTC GAGCTCGGGG    1260

CAGGGTGGTA GAACAGCTGC AACAGGGACA GGTCT    1295

a third sequence (SEQ ID NO.6):

AGGCAGACCA CATATGTGGT CGATGCC ATGGAGGCCC ATCAGTTTAT TAAGGCTCCT    57

GGCATCACTA CTGCTATTGA GCAGGCTGCT CTAGCAGCGG CCAACTCTGC CCTGGCGAAT    117

GCTGTGGTAG TTAGGCCTTT TCTCTCTCAC CAGCAGATTG AGATCCTCAT TAACCTAATG    177

CAACCTCGCC AGCTTGTTTT CCGCCCCGAG GTTTTCTGGA ATCATCCCAT CCAGCGTGTC    237

ATCCATAACG AGCTGGAGCT TTACTGCCGC GCCCGCTCCG GCCGCTGTCT TGAAATTGGC    297

GCCCATCCCC GCTCAATAAA TGATAATCCT AATGTGGTCC ACCGCTGCTT CCTCCGCCCT    357

GTTGGGCGTG ATGTTGAGCG CTGGTATACT GCTCCCACTC GCGGGCCGGC TGCTAATTGC    417

CGGCGTTCCG CGCTGCGCGG GCTTCCCGCT GCTGACCGCA CTTACTGCCT CGACGGGTTT    477

TCTGGCTGTA ACTTTCCCGC CGAGACTGGC ATCGCCCTCT ACTCCCTTCA TGATATGTCA    537

CCATCTGATG TCGCCGAGGG CATGTTCCGC CATGGTATGA CGCGGCTCTA TGCCGCCCTC    597

CATCTTCCGC CTGAGGTCCT GCTGCCCCCT GGCACATATC GCACCGCATC GTATTTGCTA    657

ATTCATGACG GTAGGCGCGT TGTGGTGACG TATGAGGGTG ATACTAGTGC TGGTTACAAC    717

CACGATGTCT CCAACTTGCG CTCCTGGATT AGAACCACCA AGGTTACCGG AGACCATCCC    777

CTCGTTATCG AGCGGGTTAG GGCCATTGGC TGCCACTTTG TTCTCTTGCT CACGGCAGCC    837

CCGGAGCCAT CACGTATGCC TTATGTTCCT TACCCCCGGT CTACCGAGGT CTATGTCCGA    897

TCGATCTTCG GCCCGGGTGG CACCCCTTCC TTATTCCCAA CCTCATGCTC CACTAAGTCG    957

ACCTTCCATG CTGTCCCTGC CCATATTTGG GACCGTCTTA TGCTGTTCGG GGCCACCTTG    1017

GATGACCAAG CCTTTTGCTG CTCCCGTTTA ATGACCTACC TTCGCGGCAT TAGCTACAAG    1077

GTCACTGTTG GTACCCTTGT GGCTAATGAA GGCTGGAATG CCTCTGAGGA CGCCCTCACA    1137

GCTGTTATCA CTGCCGCCTA CCTTACCATT TGCCACCAGC GGTATCTCCG CACCCAGGCT    1197

ATATCCAAGG GGATGCGTCG TCTGGAACGG GAGCATGCCC AGAAGTTTAT AACACGCCTC    1257

TACAGCTGGC TCTTCGAGAA GTCCGGCCGT GATTACATCC CTGGCCGTCA GTTGGAGTTC    1317

TACGCCCAGT GCAGGCGCTG GCTCTCCGCC GGCTTTCATC TTGATCCACG GGTGTTGGTT    1377

TTTGACGAGT CGGCCCCCTG CCATTGTAGG ACCGCGATCC GTAAGGCGCT CTCAAAGTTT    1437

TGCTGCTTCA TGAAGTGGCT TGGTCAGGAG TGCACCTGCT TCCTTCAGCC TGCAGAAGGC    1497

GCCGTCGGCG ACCAGGGTCA TGATAATGAA GCCTATGAGG GGTCCGATGT TGACCCTGCT    1557

GAGTCCGCCA TTAGTGACAT ATCTGGGTCC TATGTCGTCC CTGGCACTGC CCTCCAACCG    1617

CTCTACCAGG CCCTCGATCT CCCCGCTGAG ATTGTGGCTC GCGCGGGCCG GCTGACCGCC    1677

ACAGTAAAGG TCTCCCAGGT CGATGGGCGG ATCGATTGCG AGACCCTTCT TGGTAACAAA    1737

ACCTTTCGCA CGTCGTTCGT TGACGGGGCG GTCTTAGAGA CCAATGGCCC AGAGCGCCAC    1797

AATCTCTCCT TCGATGCCAG TCAGAGCACT ATGGCCGCTG GCCCTTTCAG TCTCACCTAT    1857

GCCGCCTCTG CAGCTGGGCT GGAGGTGCGC TATGTTGCTG CCGGGCTTGA CCATCGGGCG    1917

GTTTTTGCCC CCGGTGTTTC ACCCCGGTCA GCCCCCGGCG AGGTTACCGC CTTCTGCTCT    1977

GCCCTATACA GGTTTAACCG TGAGGCCCAG CGCCATTCGC TGATCGGTAA CTTATGGTTC    2037

CATCCTGAGG GACTCATTGG CCTCTTCGCC CCGTTTTCGC CCGGGCATGT TTGGGAGTCG    2097

GCTAATCCAT TCTGTGGCGA GAGCACACTT TACACCCGTA CTTGGTCGGA GGTTGATGCC    2157

GTCTCTAGTC CAGCCCGGCC TGACTTAGGT TTTATGTCTG AGCCTTCTAT ACCTAGTAGG    2217

GCCGCCACGC CTACCCTGGC GGCCCCTCTA CCCCCCCCTG CACCGGACCC TTCCCCCCCT    2277

CCCTCTGCCC CGGCGCTTGC TGAGCCGGCT TCTGGCGCTA CCGCCGGGGC CCCGGCCATA    2337

ACTCACCAGA CGGCCCGGCA CCGCCGCCTG CTCTTCACCT ACCCGGATGG CTCTAAGGTA    2397

TTCGCCGGCT CGCTGTTCGA GTCGACATGC ACGTGGCTCG TTAACGCGTC TAATGTTGAC    2457

CACCGCCCTG GCGGCGGGCT TTGCCATGCA TTTTACCAAA GGTACCCCGC CTCCTTTGAT    2517

GCTGCCTCTT TTGTGATGCG CGACGGCGCG GGCGCGTACA CACTAACCCC CCGGCCAATA    2577

ATTCACGCTG TCGCCCCTGA TTATAGGTTG GAACATAACC CAAAGAGGCT TGAGGCTGCT    2637

TATCGGGAAA CTTGCTCCCG CCTCGGCACC GCTGCATACC CGCTCCTCGG GACCGGCATA    2697

TACCAGGTGC CGATCGGCCC CAGTTTTGAC GCCTGGGAGC GGAACCACCG CCCCGGGGAT    2757

GAGTTGTACC TTCCTGAGCT TGCTGCCAGA TGGTTTGAGG CCAATAGGCC GACCCGCCCG    2817

ACTCTCACTA TAACTGAGGA TGTTGCACGG ACAGCGAATC TGGCCATCGA GCTTGACTCA    2877

GCCACAGATG TCGGCCGGGC CTGTGCCGGC TGTCGGGTCA CCCCCGGCGT TGTTCAGTAC    2937

CAGTTTACTG CAGGTGTGCC TGGATCCGGC AAGTCCCGCT CTATCACCCA AGCCGATGTG    2997

GACGTTGTCG TGGTCCCGAC GCGTGAGTTG CGTAATGCCT GGCGCCGTCG CGGCTTTGCT    3057

GCTTTTACCC CGCATACTGC CGCCAGAGTC ACCCAGGGGC GCCGGGTTGT CATTGATGAG    3117

GCTCCATCCC TCCCCCCTCA CCTGCTGCTG CTCCACATGC AGCGGGCCGC CACCGTCCAC    3177

CTTCTTGGCG ACCCGAACCA GATCCCAGCC ATCGACTTTG AGCACGCTGG GCTCGTCCCC    3237

```
GCCATCAGGC CCGACTTAGG CCCCACCTCC TGGTGGCATG TTACCCATCG CTGGCCTGCG   3297

GATGTATGCG AGCTCATCCG TGGTGCATAC CCCATGATCC AGACCACTAG CCGGGTTCTC   3357

CGTTCGTTGT TCTGGGGTGA GCCTGCCGTC GGGCAGAAAC TAGTGTTCAC CCAGGCGGCC   3417

AAGCCCGCCA ACCCCGGCTC AGTGACGGTC CACGAGGCGC AGGGCGCTAC CTACACGGAG   3477

ACCACTATTA TTGCCACAGG AGATGCCCGG GGCCTTATTC AGTCGTCTCG GGCTCATGCC   3537

ATTGTTGCTC TGACGCGCCA CACTGAGAAG TGCGTCATCA TTGACGCACC AGGCCTGCTT   3597

CGCGAGGTGG GCATCTCCGA TGCAATCGTT AATAACTTTT TCCTCGCTGG TGGCGAAATT   3657

GGTCACCAGC GCCCATCAGT TATTCCCCGT GGCAACCCTG ACGCCAATGT TGACACCCTG   3717

GCTGCCTTCC CGCCGTCTTG CCTGATTAGT GCCTTCCATC AGTTGGCTGA GGAGCTTGGC   3777

CACAGACCTG TCCCTGTTGC AGCTGTTCTA CCACCCTGCC CCGAGCTCGA ACAGGGCCTT   3837

CTCTACCTGC CCCAGGAGCT CACCAGCTGT GATAGTGTCG TAACATTTGA ATTAACAGAC   3897

ATTGTGCACT GCCGCATGGC CGCCCCGAGG CAGCGCAAGG CCGTGCTGTC CACACTCGTG   3957

GGCCGCTACG GCGGTCGCAC AAAGCTCTAC AATGCTTCCC ACTCTGATGT TCGCGACTCT   4017

CTCGCCCGTT TTATCCCGGC CATTGGCCCG GTACAGGTTA CAACTTGTGA ATTGTACGAG   4077

CTAGTGGAGG CCATGGTCGA GAAGGGCCAG GATGGCTCCG CCGTCCTTGA GCTTGATCTT   4137

TGCAACCGTG ACGTGTCCAG GATCACCTTC TTCCAGAAAG ATTGTAACAA GTTCACCACA   4197

GGTGAGACCA TTGCCCATGG TAAAGTGGGC CAGGGCATCT CGGCCTGGAG CAAGACCTTC   4257

TGCGCCCTCT TTGGCCCTTG GTTCCGCGCT ATTGAGAAGG CTATTCTGGC CCTGCTCCCT   4317

CAGGGTGTGT TTTACGGTGA TGCCTTTGAT GACACCGTCT TCTCGGCGGC TGTGGCCGCA   4377

GCAAAGGCAT CCATGGTGTT TGAGAATGAC TTTTCTGAGT TTGACTCCAC CCAGAATAAC   4437

TTTTCTCTGG GTCTAGAGTG TGCTATTATG GAGGAGTGTG GGATGCCGCA GTGGCTCATC   4497

CGCCTGTATC ACCTTATAAG GTCTGCGTGG ATCTTGCAGG CCCCGAAGGA GTCTCTGCGA   4557

GGGTTTTGGA AGAAACACTC CGGTGAGCCC GGCACTCTTC TATGGAATAC TGTCTGGAAT   4617

ATGGCCGTTA TTACCCACTG TTATGACTTC CGCGATTTTC AGGTGGCTGC CTTTAAAGGT   4677

GATGATTCGA TAGTGCTTTG CAGTGAGTAT CGTCAGAGTC CAGGAGCTGC TGTCCTGATC   4737

GCCGGCTGTG GCTTGAAGTT GAAGGTAGAT TTCCGCCCGA TCGGTTTGTA TGCAGGTGTT   4797

GTGGTGGCCC CCGGCCTTGG CGCGCTCCCT GATGTTGTGC GCTTCGCCGG CCGGCTTACC   4857

GAGAAGAATT GGGGCCCTGG CCCTGAGCGG GCGGAGCAGC TCCGCCTCGC TGTTAGTGAT   4917
```

```
TTCCTCCGCA AGCTCACGAA TGTAGCTCAG ATGTGTGTGG ATGTTGTTTC CCGTGTTTAT    4977

GGGGTTTCCC CTGGACTCGT TCATAACCTG ATTGGCATGC TACAGGCTGT TGCTGATGGC    5037

AAGGCACATT TCACTGAGTC AGTAAAACCA GTGCTCGACT TGACAAATTC AATCTTGTGT    5097

CGGGTGGAAT GA ATAACATGTC TTTTGCTGCG CCCATGGGTT CGCGACCATG            5149

CGCCCTCGGC CTATTTTGTT GCTGCTCCTC ATGTTTTTGC CTATGCTGCC CGCGCCACCG    5209

CCCGGTCAGC CGTCTGGCCG CCGTCGTGGG CGGCGCAGCG GCGGTTCCGG CGGTGGTTTC    5269

TGGGGTGACC GGGTTGATTC TCAGCCCTTC GCAATCCCCT ATATTCATCC AACCAACCCC    5329

TTCGCCCCCG ATGTCACCGC TGCGGCCGGG GCTGGACCTC GTGTTCGCCA ACCCGCCCGA    5389

CCACTCGGCT CCGCTTGGCG TGACCAGGCC CAGCGCCCCG CCGTTGCCTC ACGTCGTAGA    5449

CCTACCACAG CTGGGGCCGC GCCGCTAA CCGCGGTCGC TCCGGCCCAT GACACCCCGC      5507

CAGTGCCTGA TGTCGACTCC CGCGGCGCCA TCTTGCGCCG GCAGTATAAC CTATCAACAT    5567

CTCCCCTTAC CTCTTCCGTG GCCACCGGCA CTAACGTGGT TCTTTATGCC GCCCCTCTTA    5627

GTCCGCTTTT ACCCCTTCAG GACGGCACCA ATACCCATAT AATGGCCACG GAAGCTTCTA    5687

ATTATGCCCA GTACCGGGTT GCCCGTGCCA CAATCCGTTA CCGCCCGCTG GTCCCCAATG    5747

CTGTCGGCGG TTACGCCATC TCCATCTCAT TCTGGCCACA GACCACCACC ACCCCGACGT    5807

CCGTTGATAT GAATTCAATA ACCTCGACGG ATGTTCGTAT TTTAGTCCAG CCCGGCATAG    5867

CCTCTGAGCT TGTGATCCCA AGTGAGCGCC TACACTATCG TAACCAAGGC TGGCGCTCCG    5927

TCGAGACCTC TGGGGTGGCT GAGGAGGAGG CTACCTCTGG TCTTGTTATG CTTTGCATAC    5987

ATGGCTCACT CGTAAATTCC TATACTAATA CACCCTATAC CGGTGCCCTC GGGCTGTTGG    6047

ACTTTGCCCT TGAGCTTGAG TTTCGCAACC TTACCCCCGG TAACACCAAT ACGCGGGTCT    6107

CCCGTTATTC CAGCACTGCT CGCCACCGCC TTCGTCGCGG TGCGGACGGG ACTGCCGAGC    6167

TCACCACCAC GGCTGCTACC CGCTTTATGA AGGACCTCTA TTTTACTAGT ACTAATGGTG    6227

TCGGTGAGAT CGGCCGCGGG ATAGCCCTCA CCCTGTTCAA CCTTGCTGAC ACTCTGCTTG    6287

GCGGCCTGCC GACAGAATTG ATTTCGTCGG CTGGTGGCCA GCTGTTCTAC TCCCGTCCCG    6347

TTGTCTCAGC CAATGGCGAG CCGACTGTTA AGTTGTATAC ATCTGTAGAG AATGCTCAGC    6407

AGGATAAGGG TATTGCAATC CCGCATGACA TTGACCTCGG AGAATCTCGT GTGGTTATTC    6467

AGGATTATGA TAACCAACAT GAACAAGATC GGCCGACGCC TTCTCCAGCC CCATCGCGCC    6527

CTTTCTCTGT CCTTCGAGCT AATGATGTGC TTTGGCTCTC TCTCACCGCT GCCGAGTATG    6587
```

```
ACCAGTCCAC TTATGGCTCT TCGACTGGCC CAGTTTATGT TTCTGACTCT GTGACCTTGG    6647

TTAATGTTGG GACCGGCGCG CAGGCCGTTG CCCGGTCGCT CGATTGGACC AAGGTCACAC    6707

TTGACGGTCG CCCCCTCTCC ACCATCCAGC AGTACTCGAA GACCTTCTTT GTCCTGCCGC    6767

TCCGCGGTAA GCTCTCTTTC TGGGAGGCAG GCACAACTAA AGCCGGGTAC CCTTATAATT    6827

ATAACACCAC TGCTAGCGAC CAACTGCTTG TCGAGAATGC CGCCGGGCAC CGGGTCGCTA    6887

TTTCCACTTA CACGACTAGC CTGGGTGCTG GTCCCGTCTC CATTTCTGCG GTTGCCGTTT    6947

TAGCCCCCCA CTCTGCGCTA GCATTGCTTG AGGATACCTT GGACTACCCT GCCCGCGCCC    7007

ATACTTTTGA TGATTTCTGC CCAGAGTGCC GCCCCCTTGG CCTTCAGGGC TGCGCTTTCC    7067

AGTCTACTGT CGCTGAGCTT CAGCGCCTTA AGATGAAGGT GGGTAAAAACT CGGGAGTTGT    7127

AG TTTATTTGCT TGTGCCCCGC TTCTTTCTGT TGCTTATTTC TCATTTCTGC             7179

GTTCCGCGCT CCCTGA                                                    7195
```

a fourth sequence (SEQ ID NO.10):
```
GCCATGGAGG CCCACCAGTT CATTAAGGCT CCTGGCATCA CTACTGCTAT TGAGCAAGCA      60

GCTCTAGCAG CGGCCAACTG CGCCCTTGCG AATGCTGTGG TGGTCCGGCC TTTCCTTTCC     120

CATCAGCAGG TTGAGATCCT TATAAATCTC ATGCAACCTC GGCAGCTGGT GTTTCGTCCT     180

GAGGTTTTTT GGAATCACGC TGATTCAACGT GTTATACATA ATGAGCTTGA GCAGTATTGC     240

CGTGCTCGCT CGGGTCGCTG CCTTGAGATT GGAGCCCACC CACGCTCCAT TAATGATAAT     300

CCTAATGTCC TCCATCGCTG CTTTCTCCAC CGCGTCGGCC GGGATGTTCA GCGCTGGTAC     360

ACAGCCCCGA CTAGGGGACC TGCGGCGAAC TGTCGCCGCT CGGCACTTCG TGGTCTGCCA     420

CCAGCCGACC GCACTTACTG TTTTGATGGC TTTGCCGGCT GCCGTTTTGC CGCCGAGACT     480

GGTGTGGCTC TCTATTCTCT CCATGACTTG CAGCCGGCTG ATGTTGCCGA GGCGATGGCT     540

CGCCACGGCA TGACCCGCCT TTATGCAGCT TTCCAGTTGC CTCCAGAGGT GCTCCTGCCT     600

CCTGGCACCT ACCGGACATC ATCCTACTTG CTGATCCACG ATGGTAAGCG CGCGGTTGTC     660

ACTTATGAGG GTGACACTAG CGCCGGTTAC AATCATGATG TTGCCACCCT CCGCACATGG     720

ATCAGGACAA CTAAGGTTGT GGGTGAACAC CCTTTGGTGA TCGAGCGGGT GCGGGGTATT     780

GGCTGTCACT TTGTGTTGTT GATCACTGCG GCCCCTGAGC CTCCCCGAT GCCCTACGTT     840

CCTTACCCGC GTTCGACGGA GGTCTATGTC CGGTCTATCT TTGGGCCCGG CGGGTCCCCG     900

TCGCTGTTCC CGACCGCTTG TGCTGTCAAG TCCACTTTTC ACGCCGTCCC CACGCACATC     960
```

20309587
040491          89.

TGGGACCGTC TCATGCTCTT TGGGGCCACC CTCGACGACC AGGCCTTTTG CTGCTCCAGG    1020

CTTATGACGT ACCTTCGTGG CATTAGCTAT AAGGTAACTG TGGGTGCCCT GGTCGCTAAT    1080

GAAGGCTGGA ATGCCACCGA GGATGCGCTC ACTGCAGTTA TTACGGCGGC TTACCTCACA    1140

ATATGTCATC AGCGTTATTT GCGGACCCAG GCGATTTCTA AGGGCATGCG CCGGCTTGAG    1200

CTTGAACATG CTCAGAAATT TATTTCACGC CTCTACAGCT GGCTATTTGA GAAGTCAGGT    1260

CGTGATTACA TCCCAGGCCG CCAGCTGCAG TTCTACGCTC AGTGCCGCCG CTGGTTATCT    1320

GCCGGGTTCC ATCTCGACCC CCGCACCTTA GTTTTTGATG AGTCAGTGCC TTGTAGCTGC    1380

CGAACCACCA TCCGGCGGAT CGCTGGAAAA TTTTGCTGTT TTATGAAGTG GCTCGGTCAG    1440

GAGTGTTCTT GTTTCCTCCA GCCCGCCGAG GGGCTGGCGG GCGACCAAGG TCATGACAAT    1500

GAGGCCTATG AAGGCTCTGA TGTTGATACT GCTGAGCCTG CCACCCTAGA CATTACAGGC    1560

TCATACATCG TGGATGGTCG GTCTCTGCAA ACTGTCTATC AAGCTCTCGA CCTGCCAGCT    1620

GACCTGGTAG CTCGCGCAGC CCGACTGTCT GCTACAGTTA CTGTTACTGA AACCTCTGGC    1680

CGTCTGGATT GCCAAACAAT GATCGGCAAT AAGACTTTTC TCACTACCTT TGTTGATGGG    1740

GCACGCCTTG AGGTTAACGG GCCTGAGCAG CTTAACCTCT CTTTTGACAG CCAGCAGTGT    1800

AGTATGGCAG CCGGCCCGTT TTGCCTCACC TATGCTGCCG TAGATGGCGG GCTGGAAGTT    1860

CATTTTTCCA CCGCTGGCCT CGAGAGCCGT GTTGTTTTCC CCCCTGGTAA TGCCCCGACT    1920

GCCCCGCCGA GTGAGGTCAC GGCTTCTGC TCAGCTCTTT ATAGGCACAA CCGGCAGAGC    1980

CAGCGCCAGT CGGTTATTGG TAGTTTGTGG CTGCACCCTG AAGGTTTGCT CGGCCTGTTC    2040

CCGCCCTTTT CACCCGGGCA TGAGTGGCGG TCTGCTAACC CATTTTGCGG CGAGAGCACG    2100

CTCTACACCC GCACTTGGTC CACAATTACA GACACACCCT TAACTGTCGG GCTAATTTCC    2160

GGTCATTTGG ATGCTGCTCC CCACTCGGGG GGGCCACCTG CTACTGCCAC AGGCCCTGCT    2220

GTAGGCTCGT CTGACTCTCC AGACCCTGAC CCGCTACCTG ATGTTACAGA TGGCTCACGC    2280

CCCTCTGGGG CCCGTCCGGC TGGCCCCAAC CCGAATGGCG TTCCGCAGCG CCGCTTACTA    2340

CACACCTACC CTGACGGCGC TAAGATCTAT GTCGGCTCCA TTTTCGAGTC TGAGTGCACC    2400

TGGCTTGTCA ACGCATCTAA CGCCGGCCAC CGCCCTGGTG GCGGGCTTTG TCATGCTTTT    2460

TTTCAGCGTT ACCCTGATTC GTTTGACGCC ACCAAGTTTG TGATGCGTGA TGGTCTTGCC    2520

GCGTATACCC TTACACCCCG GCCGATCATT CATGCGGTGG CCCCGGACTA TCGATTGGAA    2580

CATAACCCCA AGAGGCTCGA GGCTGCCTAC CGCGAGACTT GCGCCCGCCG AGGCACTGCT    2640

```
GCCTATCCAC TCTTAGGCGC TGGCATTTAC CAGGTGCCTG TTAGTTTGAG TTTTGATGCC    2700

TGGGAGCGGA ACCACCGCCC GTTTGACGAG CTTTACCTAA CAGAGCTGGC GGCTCGGTGG    2760

TTTGAATCCA ACCGCCCCGG TCAGCCCACG TTGAACATAA CTGAGGATAC CGCCCGTGCG    2820

GCCAACCTGG CCCTGGAGCT TGACTCCGGG AGTGAAGTAG GCCGCGCATG TGCCGGGTGT    2880

AAAGTCGAGC CTGGCGTTGT GCGGTATCAG TTTACAGCCG GTGTCCCCGG CTCTGGCAAG    2940

TCAAAGTCCG TGCAACAGGC GGATGTGGAT GTTGTTGTTG TGCCCACTCG CGAGCTTCGG    3000

AACGCTTGGC GGCGCCGGGG CTTTGCGGCA TTCACTCCGC ACACTGCGGC CCGTGTCACT    3060

AGCGGCCGTA GGGTTGTCAT TGATGAGGCC CCTTCGCTCC CCCCACACTT GCTGCTTTTA    3120

CATATGCAGC GTGCTGCATC TGTGCACCTC CTTGGGGACC CGAATCAGAT CCCCGCCATA    3180

GATTTTGAGC ACACCGGTCT GATTCCAGCA ATACGGCCGG AGTTGGTCCC GACTTCATGG    3240

TGGCATGTCA CCCACCGTTG CCCTGCAGAT GTCTGTGAGT TAGTCCGTGG TGCTTACCCT    3300

AAAATCCAGA CTACAAGTAA GGTGCTCCGT TCCCTTTTCT GGGGAGAGCC AGCTGTCGGC    3360

CAGAAGCTAG TGTTCACACA GGGTGCTAAG GCCGCGCACC CCGGATCTAT AACGGTCCAT    3420

GAGGCCCAGG GTGCCACTTT TACCACTACA ACTATAATTG CAACTGCAGA TGCCCGTGGC    3480

CTCATACAGT CCTCCCGGGC TGACGCTATA GTTGCTCTCA CTAGGCATAC TGAAAAATGT    3540

GTTATACTTG ACTCTCCCGG CCTGTTGCGT GAGGTGGGTA TCTCAGATGC CATTGTTAAT    3600

AATTTCTTCC TTTCGGGTGG CGAGGTTGGT CACCAGAGAC CATCGGTCAT TCCGCGAGGC    3660

AACCCTGACC GCAATGTTGA CGTGCTTGCG GCGTTTCCAC CTTCATGCCA AATAAGCGCC    3720

TTCCATCAGC TTGCTGAGGA GCTGGGCCAC CGGCCGGCGC CGGTGGCGGC TGTGCTACCT    3780

CCCTGCCCTG AGCTTGAGCA GGGCCTTCTC TATCTGCCAC AGGAGCTAGC CTCCTGTGAC    3840

AGTGTTGTGA CATTTGAGCT AACTGACATT GTGCACTGCC GCATGGCGGC CCCTAGCCAA    3900

AGGAAAGCTG TTTTGTCCAC GCTGGTAGGC CGGTATGGCA GACGCACAAG GCTTTATGAT    3960

GCGGGTCACA CCGATGTCCG CGCCTCCCTT GCGCGCTTTA TTCCCACTCT CGGGCGGGTT    4020

ACTGCCACCA CCTGTGAACT CTTTGAGCTT GTAGAGGCGA TGGTGGAGAA GGGCCAAGAC    4080

GGTTCAGCCG TCCTCGAGTT GGATTTGTGC AGCCGAGATG TCTCCGGCAT AACCTTTTTC    4140

CAGAAGGATT GTAACAAGTT CACGACCGGC GAGACAATTG CGCATGGCAA AGTCGGTCAG    4200

GGTATCTTCC GCTGGAGTAA GACGTTTTGT GCCCTGTTTG GCCCCTGGTT CCGTGCGATT    4260

GAGAAGGCTA TTCTATCCCT TTTACCACAA GCTGTGTTCT ACGGGGATGC TTATGACGAC    4320
```

```
TCAGTATTCT CTGCTGCCGT GGCTGGCGCC AGCCATGCCA TGGTGTTTGA AAATGATTTT    4380

TCTGAGTTTG ACTCGACTCA GAATAACTTT TCCCTAGGTC TTGAGTGCGC CATTATGGAA    4440

GAGTGTGGTA TGCCCCAGTG GCTTGTCAGG TTGTACCATG CCGTCCGGTC GGCGTGGATC    4500

CTGCAGGCCC CAAAAGAGTC TTTGAGAGGG TTCTGGAAGA AGCATTCTGG TGAGCCGGGC    4560

AGCTTGCTCT GGAATACGGT GTGGAACATG GCAATCATTG CCCATTGCTA TGAGTTCCGG    4620

GACCTCCAGG TTGCCGCCTT CAAGGGCGAC GACTCGGTCG TCCTCTGTAG TGAATACCGC    4680

CAGAGCCCAG GCGCCGGTTC GCTTATAGCA GGCTGTGGTT TGAAGTTGAA GGCTGACTTC    4740

CGGCCGATTG GGCTGTATGC CGGGGTTGTC GTCGCCCCGG GGCTCGGGGC CCTACCCGAT    4800

GTCGTTCGAT TCGCCGGACG GCTTTCGGAG AAGAACTGGG GGCCTGATCC GGAGCGGGCA    4860

GAGCAGCTCC GCCTCGCCGT GCAGGATTTC CTCCGTAGGT TAACGAATGT GGCCCAGATT    4920

TGTGTTGAGG TGGTGTCTAG AGTTTACGGG GTTTCCCCGG GTCTGGTTCA TAACCTGATA    4980

GGCATGCTCC AGACTATTGG TGATGGTAAG GCGCATTTTA CAGAGTCTGT TAAGCCTATA    5040

CTTGACCTTA CACACTCAAT TATGCACCGG TCTGAATGAA TAACATGTGG TTTGCTGCGC    5100

CCATGGGTTC GCCACCATGC GCCCTAGGCC TCTTTTGCTG TTGTTCCTCT TGTTTCTGCC    5160

TATGTTGCCC GCGCCACCGA CCGGTCAGGC GTCTGGCCGC CGTCGTGGGC GGCGCAGCGG    5220

CGGTACCGGC GGTGGTTTCT GGGGTGACCG GGTTGATTCT CAGCCCTTCG CAATCCCCTA    5280

TATTCATCCA ACCAACCCCT TTGCCCCAGA CGTTGCCGCT GCGTCCGGGT CTGGACCTCG    5340

CCTTCGCCAA CCAGCCCGGC CACTTGGCTC CACTTGGCGA GATCAGGCCC AGCGCCCCTC    5400

CGCTGCCTCC CGTCGCCGAC CTGCCACAGC CGGGGCTGCG GCGCTGACGG CTGTGGCGCC    5460

TGCCCATGAC ACCTCACCCG TCCCGGACGT TGATTCTCGC GGTGCAATTC TACGCCGCCA    5520

GTATAATTTG TCTACTTCAC CCCTGACATC CTCTGTGGCC TCTGGCACTA ATTTAGTCCT    5580

GTATGCAGCC CCCCTTAATC CGCCTCTGCC GCTGCAGGAC GGTACTAATA CTCACATTAT    5640

GGCCACAGAG GCCTCCAATT ATGCACAGTA CCGGGTTGCC CGCGCTACTA TCCGTTACCG    5700

GCCCCTAGTG CCTAATGCAG TTGGAGGCTA TGCTATATCC ATTTCTTTCT GGCCTCAAAC    5760

AACCACAACC CCTACATCTG TTGACATGAA TTCCATTACT TCCACTGATG TCAGGATTCT    5820

TGTTCAACCT GGCATAGCAT CTGAATTGGT CATCCCAAGC GAGCGCCTTC ACTACCGCAA    5880

TCAAGGTTGG CGCTCGGTTG AGACATCTGG TGTTGCTGAG GAGGAAGCCA CCTCCGGTCT    5940

TGTCATGTTA TGCATACATG GCTCTCCAGT TAACTCCTAT ACCAATACCC CTTATACCGG    6000
```

TGCCCTTGGC TTACTGGACT TTGCCTTAGA GCTTGAGTTT CGCAATCTCA CCACCTGTAA      6060

CACCAATACA CGTGTGTCCC GTTACTCCAG CACTGCTCGT CACTCCGCCC GAGGGGCCGA      6120

CGGGACTGCG GAGCTGACCA CAACTGCAGC CACCAGGTTC ATGAAAGATC TCCACTTTAC      6180

CGGCCTTAAT GGGGTAGGTG AAGTCGGCCG CGGGATAGCT CTAACATTAC TTAACCTTGC      6240

TGACACGCTC CTCGGCGGGC TCCCGACAGA ATTAATTTCG TCGGCTGGCG GGCAACTGTT      6300

TTATTCCCGC CCGGTTGTCT CAGCCAATGG CGAGCCAACC GTGAAGCTCT ATACATCAGT      6360

GGAGAATGCT CAGCAGGATA AGGGTGTTGC TATCCCCCAC GATATCGATC TTGGTGATTC      6420

GCGTGTGGTC ATTCAGGATT ATGACAACCA GCATGAGCAG GATCGGCCCA CCCCGTCGCC      6480

TGCGCCATCT CGGCCTTTTT CTGTTCTCCG AGCAAATGAT GTACTTTGGC TGTCCCTCAC      6540

TGCAGCCGAG TATGACCAGT CCACTTACGG GTCGTCAACT GGCCCGGTTT ATATCTCGGA      6600

CAGCGTGACT TTGGTGAATG TTGCGACTGG CGCGCAGGCC GTAGCCCGAT CGCTTGACTG      6660

GTCCAAAGTC ACCCTCGACG GCGGCCCCT CCCGACTGTT GAGCAATATT CCAAGACATT      6720

CTTTGTGCTC CCCCTTCGTG GCAAGCTCTC CTTTTGGGAG GCCGGCACAA CAAAAGCAGG      6780

TTATCCTTAT AATTATAATA CTACTGCTAG TGACCAGATT CTGATTGAAA ATGCTGCCGG      6840

CCATCGGGTC GCCATTTCAA CCTATACCAC CAGGCTTGGG GCCGGTCCGG TCGCCATTTC      6900

TGCGGCCGCG GTTTTGGCTC CACGCTCCGC CCTGGCTCTG CTGGAGGATA CTTTTGATTA      6960

TCCGGGGCGG GCGCACACAT TTGATGACTT CTGCCCTGAA TGCCGCGCTT TAGGCCTCCA      7020

GGGTTGTGCT TTCCAGTCAA CTGTCGCTGA GCTCCAGCGC CTTAAAGTTA AGGTGGGTAA      7080

AACTCGGGAG TTGTAGTTTA TTTGGCTGTG CCCACCTACT TATATCTGCT GATTTCCTTT      7140

ATTTCCTTTT TCTCGGTCCC GCGCTCCCTG A                                     7171

or a fifth sequence (SEQ ID NO.12):

CGGGCCCCGT ACAGGTCACA ACCTGTGAGT TGTACGAGCT AGTGGAGGCC ATGGTCGAGA        60

AAGGCCAGGA TGGCTCCGCC GTCCTTGAGC TCGATCTCTG CAACCGTGAC GTGTCCAGGA       120

TCACCTTTTT CCAGAAAGAT TGCAATAAGT TCACCACGGG AGAGACCATC GCCCATGGTA       180

AAGTGGGCCA GGGCATTTCG GCCTGGAGTA AGACCTTCTG TGCCCTTTTC GGCCCCTGGT       240

TCCGTGCTAT TGAGAAGGCT ATTCTGGCCC TGCTCCCTCA GGGTGTGTTT TATGGGGATG       300

CCTTTGATGA CACCGTCTTC TCGGCGCGTG TGGCCGCAGC AAAGGCGTCC ATGGTGTTTG       360

AGAATGACTT TTCTGAGTTT GACTCCACCC AGAATAATTT TTCCCTGGGC CTAGAGTGTG       420

CTATTATGGA GAAGTGTGGG ATGCCGAAGT GGCTCATCCG CTTGTACCAC CTTATAAGGT       480

```
CTGCGTGGAT CCTGCAGGCC CCGAAGGAGT CCCTGCGAGG GTGTTGGAAG AAACACTCCG    540

GTGAGCCCGG CACTCTTCTA TGGAATACTG TCTGGAACAT GGCCGTTATC ACCCATTGTT    600

ACGATTTCCG CGATTTGCAG GTGGCTGCCT TTAAAGGTGA TGATTCGATA GTGCTTTGCA    660

GTGAGTACCG TGAGAGTCCA GGGGCTGCTG TCCTGATTGC TGGCTGTGGC TTAAAGCTGA    720

AGGTGGGTTT CCGTCCGATT GGTTTGTATG CAGGTGTTGT GGTGACCCCC GGCCTTGGCG    780

CGCTTCCCGA CGTCGTGCGC TTGTCCGGCC GGCTTACTGA GAAGAATTGG GGCCCTGGCC    840

CTGAGCGGGC GGAGCAGCTC CGCCTTGCTG TGCG    874
```

or a sequence complementary thereto.

4. A protein which is (a) immunoreactive with antibodies present in individuals infected with enterically transmitted nonA/nonB hepatitis and (b) derived from a viral hepatitis agent whose genome contains a region which is homologous to the 1.33 kb DNA EcoRI insert present in plasmid pTZXF1(ET1.1) carried in E. coli strain BB4, and having ATCC Deposit Nno. 67717.

5. The protein of claim 4, which is encoded by a coding region within said 1.33 kb EcoRI insert.

6. A protein which is (a) immunoreactive with antibodies present in individuals infected with enterically transmitted nonA/nonB hepatitis and (b) encoded by genetic sequence 406.3-2 or 406.4-2 or a fragment thereof.

7. A method of detecting infection by enterically transmitted nonA/nonB hepatitis viral agent in a test individual, comprising:
providing a peptide antigen which is (a) immunoreactive with antibodies present in individuals infected with enterically transmitted nonA/nonB hepatitis and (b) derived from a viral hepatitis agent whose genome contains a region which is homologous to

20309587
040491                    94.

the 1.33 kb DNA EcoRI insert present in plasmid pTZKF1(ET1.1) carried in E. coli strain BB4, and having ATCC deposit no. 67717,

     reacting serum from the test individual with

5   such antigen, and

     examining the antigen for the presence of bound antibody.


    8.   The method of claim 7, wherein the serum

10  antibody is an IgM or IgG antibody, or a mixture of both, the antigen provided is attached to a support, said reacting includes contacting such serum with the support and said examining includes reacting the

15  support and bound serum antibody with a reporter-labeled anti-human antibody.


    9.   A kit for ascertaining the presence of serum antibodies which are diagnostic of enterically transmitted nonA/nonB hepatitis infection, comprising

20     a support with surface-bound recombinant peptide antigen which is (a) immunoreactive with antibodies present in individuals infected with enterically transmitted nonA/nonB viral hepatitis agent and (b) derived from a viral hepatitis agent

25  whose genome contains a region which is homologous to the 1.33 kb DNA EcoRI insert present in plasmid pTZKF1(ET1.1) carried in E. coli strain BB4, and having ATCC deposit no. 67717, and

     a reporter-labeled anti-human antibody.

30

    10.  A DNA fragment derived from an enterically transmitted nonA/nonB viral hepatitis agent whose genome contains a region which is homologous to the 1.33 kb DNA EcoRI insert present in plasmid

35  pTZKF1(ET1.1) carried in E. coli strain BB4 and having ATCC deposit no. 67717.

11.  The fragment of claim 10, which is derived from said 1.33 kb EcoRI insert.

12.  A DNA molecule comprising genetic sequence 406.3-2 or 406.4-2 or a fragment thereof, wherein said fragment comprises at least 12 consecutive nucleotides.

13.  A DNA fragment derived from an enterically transmitted nonA/nonB viral hepatitis agent whose genome contains a region which is homologous to a DNA fragment within a first sequence (SEQ ID NO.1):

```
AGACCTGTCC CTGTTGCAGC TGTTCTACCA CCCTGCCCCG AGCTCGAACA GGGCCTTCTC    60

TACCTGCCCC AGGAGCTCAC CACCTGTGAT AGTGTCGTAA CATTTGAATT AACAGACATT   120

GTGCACTGCC GCATGGCCGC CCCGAGCCAG CGCAAGGCCG TGCTGTCCAC ACTCGTGGGC   180

CGCTACGGCG GTCGCACAAA GCTCTACAAT GCTTCCCACT CTGATGTTCG CGACTCTCTC   240

GCCCGTTTTA TCCCGGCCAT TGGCGCCGTA GAGGTTACAA CTTGTGAATT GTACGAGCTA   300

GTGGAGGCCA TGGTCGAGAA GGGCCAGGAT GGCTCCGCCG TCCTTGAGCT TGATCTTTGC   360

AACCGTGACG TGTCCAGGAT CACCTTCTTC CAGAAAGATT GTAACAAGTT CACCACAGGT   420

GAGACCATTG CCCATGGTAA AGTGGGCCAG GGCATCTCGG CCTGGAGCAA GACCTTCTGC   480

GCCCTCTTTG GCCCTTGGTT CCGCGCTATT GAGAAGGCTA TTCTGGCCCT GCTCCCTCAG   540

GGTGTGTTTT ACGGTGATGC CTTTGATGAC ACCGTCTTCT CGGCGGCTGT GGCCGCAGCA   600

AAGGCATCCA TGGTGTTTGA GAATGACTTT TCTGAGTTTG ACTCCACCCA GAATAACTTT   660

TCTCTGGGTC TAGAGTGTGC TATTATGGAG GAGTGTGGGA TGCCGCAGTG GCTCATCCGC   720

CTGTATCACC TTATAAGGTC TGCGTGGATC TTGCAGGCCC CGAAGGAGTC TCTGCGAGGG   780

TTTTGGAAGA AACACTCCGG TGAGCCCGGC ACTCTTCTAT GGAATACTGT CTGGAATATG   840

GCCGTTATTA CCCACTGTTA TGACTTCCGC GATTTTCAGG TGGCTGCCTT TAAAGGTGAT   900

GATTCGATAG TGCTTTGCAG TGAGTATCGT CAGAGTCCAG GAGCTGCTGT CCTGATCGCC   960

GGCTGTGGCT TGAAGTTGAA GGTAGATTTC CGCCCGATCG GTTTGTATGC AGGTGTTGTG  1020

GTGGCCCCCG GCCTTGGCGC GCTCCCTGAT GTTGTGCGCT TCGCCGGCCG GCTTACCGAG  1080

AAGAATTGGG GCCCTGGCCC TGAGCGGGCG GAGCAGCTCC GCCTCGCTGT TAGTGATTTC  1140
```

96

CTCCGCAAGC TCACGAATGT AGCTCAGATG TGTGTGGATG TTGTTTCCCG TGTTTATGGG    1200

GTTTCCCCTG GACTCGTTCA TAACCTGATT GGCATGCTAC AGGCTGTTGC TGATGGCAAG    1260

GCACATTTCA CTGAGTCAGT AAAACCAGTG CTCGA    1295

a second sequence (SEQ ID NO.5):

TCGAGCACTG GTTTTACTGA CTCAGTGAAA TGTGCCTTGC CATCAGCAAC AGCCTGTAGC    60

ATGCCAATCA GGTTATGAAG GAGTCCAGGG GAAACCCCAT AAACACGGGA AACAACATCC    120

ACACACATCT GAGCTACATT CGTGAGCTTG CGGAGGAAAT CACTAACAGC GAGGCGGAGC    180

TGCTCCGCCC GCTCAGGGCC AGGGCCCCAA TTCTTCTCGG TAAGCCGGCC GGCGAAGCGC    240

ACAACATCAG GGAGCGCGCC AAGGCCGGGG GCCACCACAA CACCTGCATA CAAACCGATC    300

GGGCGGAAAT CTACCTTCAA CTTCAAGCCA CAGCCGGCGA TCAGGACAGC AGCTCCTGGA    360

CTCTGACGAT ACTCACTGCA AAGCACTATC GAATCATCAC CTTTAAAGGC AGCCACCTGA    420

AAATCGCGGA AGTCATAACA GTGGGTAATA ACGGCCATAT TCCAGACAGT ATTCCATAGA    480

AGAGTGCCGG GCTCACCGGA GTGTTTCTTC CAAAACCCTC GCAGAGACTC CTTCGGGGCC    540

TGCAAGATCC ACGCAGACCT TATAAGGTGA TACAGGCGGA TGAGCCACTG CGGCATCCCA    600

CACTCCTCCA TAATAGCACA CTCTAGACCC AGAGAAAAGT TATTCTGGGT GGAGTCAAAC    660

TCAGAAAAGT CATTCTCAAA CACCATGGAT GCCTTTGCTG CGGCCACAGC CGCCGAGAAG    720

ACGGTGTCAT CAAAGGCATC ACCGTAAAAC ACACCCTGAG GGAGCAGGGC CAGAATAGCC    780

TTCTCAATAG CGCGGAACCA AGGGCCAAAG AGGGCGCAGA AGGTCTTGCT CCAGGCCGAG    840

ATGCCCTGGC CCACTTTACC ATGGGCAATG GTCTCACCTG TGGTGAACTT GTTACAATCT    900

TTCTGGAAGA AGGTGATCCT GGACACGTCA CGGTTGCAAA GATCAAGCTC AAGGACGGCG    960

GAGCCATCCT GGCCCTTCTC GACCATGGCC TCCACTAGCT CGTACAATTC ACAAGTTGTA    1020

ACCTGTACGG GGCCAATGGC CGGGATAAAA CGGGCGAGAG AGTCGCGAAC ATCAGAGTGG    1080

GAAGCATTGT AGAGCTTTGT GCGACCGCCG TAGCGGCCCA CGAGTGTGGA CAGCACGGCC    1140

TTGCGCTGGC TCGGGGCGGC CATGCGGCAG TGCACAATGT CTGTTAATTC AAATGTTACG    1200

ACACTATCAC AGGTGGTGAG CTCCTGGGGC AGGTAGAGAA GGCCCTGTTC GAGCTCGGGG    1260

CAGGGTGGTA GAACAGCTGC AACAGGGACA GGTCT    1295

a third sequence (SEQ ID NO.6):

AGGCAGACCA CATATGTGGT CGATGCC ATGGAGGCCC ATCAGTTTAT TAAGGCTCCT    57

GGCATCACTA CTGCTATTGA GCAGGCTGCT CTAGCAGCGG CCAACTCTGC CCTGGCGAAT    117

97

```
GCTGTGGTAG TTAGGCCTTT TCTCTCTCAC CAGCAGATTG AGATCCTCAT TAACCTAATG    177

CAACCTCGCC AGCTTGTTTT CCGCCCCGAG GTTTTCTGGA ATCATCCCAT CCAGCGTGTC    237

ATCCATAACG AGCTGGAGCT TTACTGCCGC GCCCGCTCCG GCCGCTGTCT TGAAATTGGC    297

GCCCATCCCC GCTCAATAAA TGATAATCCT AATGTGGTCC ACCGCTGCTT CCTCCGCCCT    357

GTTGGGCGTG ATGTTCAGCG CTGGTATACT GCTCCCACTC GCGGGCCGGC TGCTAATTGC    417

CGGCGTTCCG CGCTGCGCGG GCTTCCCGCT GCTGACCGCA CTTACTGCCT CGACGGGTTT    477

TCTGGCTGTA ACTTTCCCGC CGAGACTGGC ATCGCCCTCT ACTCCCTTCA TGATATGTCA    537

CCATCTGATG TCGCCGAGGC CATGTTCCGC CATGGTATGA CGCGGCTCTA TGCCGCCCTC    597

CATCTTCCGC CTGAGGTCCT GCTGCCCCT GGCACATATC GCACCGCATC GTATTTGCTA    657

ATTCATGACG GTAGGCGCGT TGTGGTGACG TATGAGGGTG ATACTAGTGC TGGTTACAAC    717

CACGATGTCT CCAACTTGCG CTCCTGGATT AGAACCACCA AGGTTACCGG AGACCATCCC    777

CTCGTTATCG AGCGGGTTAG GGCCATTGGC TGCCACTTTG TTCTCTTGCT CACGGCAGCC    837

CCGGAGCCAT CACCTATGCC TTATGTTCCT TACCCCCGGT CTACCGAGGT CTATGTCCGA    897

TCGATCTTCG GCCCGGGTGG CACCCCTTCC TTATTCCCAA CCTCATGCTC CACTAAGTCG    957

ACCTTCCATG CTGTCCCTGC CCATATTTGG GACCGTCTTA TGCTGTTCGG GGCCACCTTG    1017

GATGACCAAG CCTTTTGCTG CTCCCGTTTA ATGACCTACC TTCGCGGCAT TAGCTACAAG    1077

GTCACTGTTG GTACCCTTGT GGCTAATGAA GGCTGGAATG CCTCTGAGGA CGCCCTCACA    1137

GCTGTTATCA CTGCCGCCTA CCTTACCATT TGCCACCAGC GGTATCTCCG CACCCAGGCT    1197

ATATCCAAGG GGATGCGTCG TCTGGAACGG GAGCATGCCC AGAAGTTTAT AACACGCCTC    1257

TACAGCTGGC TCTTCGAGAA GTCCGGCCGT GATTACATCC CTGGCCGTCA GTTGGAGTTC    1317

TACGCCCAGT GCAGGCGCTG GCTCTCCGCC GGCTTTCATC TTGATCCACG GGTGTTGGTT    1377

TTTGACGAGT CGGCCCCCTG CCATTGTAGG ACCGCGATCC GTAAGGCGCT CTCAAAGTTT    1437

TGCTGCTTCA TGAAGTGGCT TGGTCAGGAG TGCACCTGCT TCCTTCAGCC TGCAGAAGGC    1497

GCCGTCGGCG ACCAGGGTCA TGATAATGAA GCCTATGAGG GGTCCGATGT TGACCCTGCT    1557

GAGTCCGCCA TTAGTGACAT ATCTGGGTCC TATGTCGTCC CTGGCACTGC CCTCCAACCG    1617

CTCTACCAGG CCCTCGATCT CCCCGCTGAG ATTGTGGCTC GCGCGGGCCG GCTGACCGCC    1677

ACAGTAAAGG TCTCCCAGGT CGATGGGCGG ATCGATTGCG AGACCCTTCT TGGTAACAAA    1737

ACCTTTCGCA CGTCGTTCGT TGACGGGGCG GTCTTAGAGA CCAATGGCCC AGAGCGCCAC    1797
```

98

99

```
ACCACTATTA TTGGCACAGC AGATGCCCGG GGCCTTATTC AGTCGTCTCG GGCTCATGCC    3537

ATTGTTGCTC TGACGGGCCA CACTGAGAAG TGCGTCATCA TTGACGCACC AGGCCTGCTT    3597

CGCGAGGTGG GCATCTCCGA TGCAATCGTT AATAACTTTT TCCTCGCTGG TGGCGAAATT    3657

GGTCACCAGC GCCCATCAGT TATTCCCCGT GGCAACCCTG ACGCCAATGT TGACACCCTG    3717

GCTGCCTTCC CGCCGTCTTG CCAGATTAGT GCCTTCCATC AGTTGGCTGA GGAGCTTGGC    3777

CACAGACCTG TCCCTGTTGC AGCTGTTCTA CCACCCTGCC CCGAGCTCGA ACAGGGCCTT    3837

CTCTACCTGC CCCAGGAGCT CACCACCTGT GATAGTGTCG TAACATTTGA ATTAACAGAC    3897

ATTGTGCACT GCCGCATGGC CGCCCCGAGC CAGCGCAAGG CCGTGCTGTC CACACTCGTG    3957

GGCCGCTACG GCGGTCGCAC AAAGCTCTAC AATGCTTCCC ACTCTGATGT TCGCGACTCT    4017

CTCGCCCGTT TTATCCCGGC CATTGGCCCC GTACAGGTTA CAACTTGTGA ATTGTACGAG    4077

CTAGTGGAGG CCATGGTCGA GAAGGGCCAG GATGGCTCCG CCGTCCTTGA GCTTGATCTT    4137

TGCAACCGTG ACGTGTCCAG GATCACCTTC TTCCAGAAAG ATTGTAACAA GTTCACCACA    4197

GGTGAGACCA TTGCCCATGG TAAAGTGGGC CAGGGCATCT CGGCCTGGAG CAAGACCTTC    4257

TGCGCCCTCT TTGGCCCTTG GTTCCGCGCT ATTGAGAAGG CTATTCTGGC CCTGCTCCCT    4317

CAGGGTGTGT TTTACGGTGA TGCCTTTGAT GACACCGTCT TCTCGGCGGC TGTGGCCGCA    4377

GCAAAGGCAT CCATGGTGTT TGAGAATGAC TTTTCTGAGT TTGACTCCAC CCAGAATAAC    4437

TTTTCTCTGG GTCTAGAGTG TGCTATTATG GAGGAGTGTG GGATGCCGCA GTGGCTCATC    4497

CGCCTGTATC ACCTTATAAG GTCTGCGTGG ATCTTGCAGG CCCCGAAGGA GTCTCTGCGA    4557

GGGTTTTGGA AGAAACACTC CGGTGAGCCC GGCACTCTTC TATGGAATAC TGTCTGGAAT    4617

ATGGCCGTTA TTACCCACTG TTATGACTTC CGCGATTTTC AGGTGGCTGC CTTTAAAGGT    4677

GATGATTCGA TAGTGCTTTG CAGTGAGTAT CGTCAGAGTC CAGGAGCTGC TGTCCTGATC    4737

GCCGGCTGTG GCTTGAAGTT GAAGGTAGAT TTCCGCCCGA TCGGTTTGTA TGCAGGTGTT    4797

GTGGTGGCCC CCGGCCTTGG CGCGCTCCCT GATGTTGTGC GCTTCGCCGG CCGGCTTACC    4857

GAGAAGAATT GGGGCCCTGG CCCTGAGCGG GCGGAGCAGC TCCGCCTCGC TGTTAGTGAT    4917

TTCCTCCGCA AGCTCACGAA TGTAGCTCAG ATGTGTGTGG ATGTTGTTTC CCGTGTTTAT    4977

GGGGTTTCCC CTGGACTCGT TCATAACCTG ATTGGCATGC TACAGGCTGT TGCTGATGGC    5037

AAGGCACATT TCACTGAGTC AGTAAAACCA GTGCTCGACT TGACAAATTC AATCTTGTGT    5097

CGGGTGGAAT GA ATAACATGTC TTTTGCTGCG CCCATGGGTT CGCGACCATG             5149
```

100

```
                              ●                              ●

          CGCCCTCGGC CTATTTTGTT GCTGCTCCTC ATGTTTTTGC CTATGCTGCC CGCGCCACCG   5209

          CCCGGTCAGC CGTCTGGCCG CCGTCGTGGG CGGCGCAGCG GCGGTTCCGG CGGTGGTTTC   5269

    5     TGGGGTGACC GGGTTGATTC TCAGCCCTTC GCAATCCCCT ATATTCATCC AACCAACCCC   5329

          TTCGCCCCCG ATGTCACCGC TGCGGCCGGG GCTGGACCTC GTGTTCGCCA ACCCGCCCGA   5389

    10    CCACTCGGCT CCGCTTGGCG TGACCAGGCC CAGCGCCCCG CCGTTGCCTC ACGTCGTAGA   5449

          CCTACCACAG CTGGGGCCGC GCCGCTAA CCGCGGTCGC TCCGGCCCAT GACACCCCGC      5507

          CAGTGCCTGA TGTCGACTCC CGCGGCGCCA TCTTGCGCCG GCAGTATAAC CTATCAACAT   5567

    15    CTCCCCTTAC CTCTTCCGTG GCCACCGGCA CTAACCTGGT TCTTTATGCC GCCCCTCTTA   5627

          GTCCGCTTTT ACCCCTTCAG GACGGCAGCA ATACCCATAT AATGGCCACG GAAGCTTCTA   5687

    20    ATTATGCCCA GTACCGGGTT GCCCGTGCCA CAATCCGTTA CCGCCCGCTG GTCCCCAATG   5747

          CTGTCGGCGG TTACGCCATC TCCATCTCAT TCTGGCCACA GACCACCACC ACCCCGACGT   5807

          CCGTTGATAT GAATTCAATA ACCTCGACGG ATGTTCGTAT TTTAGTCCAG CCCGGCATAG   5867

    25    CCTCTGAGCT TGTGATCCCA AGTGAGCGCC TACACTATCG TAACCAAGGC TGGCGCTCCG   5927

          TCGAGACCTC TGGGGTGGCT GAGGAGGAGG CTACCTCTGG TGTTGTTATG CTTTGCATAC   5987

    30    ATGGCTCACT CGTAAATTCC TATACTAATA CACCCTATAC CGGTGCCCTC GGGCTGTTGG   6047

          ACTTTGCCCT TGAGCTTGAG TTTCGCAACC TTACCCCCGG TAACACCAAT ACGCGGGTCT   6107

          CCCGTTATTC CAGCACTGCT CGCCACCGCC TTCGTCGCGG TGCGGACGGG ACTGCCGAGC   6167

    35    TCACCACCAC GGCTGCTACC CGCTTTATGA AGGACCTCTA TTTTACTAGT ACTAATGGTG   6227

          TCGGTGAGAT CGGCCGCGGG ATAGCCCTCA CCCTGTTCAA CCTTGCTGAC ACTCTGCTTG   6287

    40    GCGGCCTGCC GACAGAATTG ATTTCGTCGG CTGGTGGCCA GCTGTTCTAC TCCCGTCCCG   6347

          TTGTCTCAGC CAATGGCGAG CCGACTGTTA AGTTGTATAC ATCTGTAGAG AATGCTCAGC   6407

    45    AGGATAAGGG TATTGCAATC CCGCATGACA TTGACCTCGG AGAATCTCGT GTGGTTATTC   6467

          AGGATTATGA TAACCAACAT GAACAAGATC GGCCGACGCC TTCTCCAGCC CCATCGCGCC   6527

          CTTTCTCTGT CCTTCGAGCT AATGATGTGC TTTGGCTCTC TCTCACCGCT GCCGAGTATG   6587

    50    ACCAGTCCAC TTATGGCTCT TCGACTGGCC CAGTTTATGT TTCTGACTCT GTGACCTTGG   6647

          TTAATGTTGC GACCGGCGCG CAGGCCGTTG CCCGGTCGCT CGATTGGACC AAGGTCACAC   6707

          TTGACGGTCG CCCCCTCTCC ACCATCCAGC AGTACTCGAA GACCTTCTTT GTCCTGCCGC   6767

    55    TCCGCGGTAA GCTCTCTTTC TGGGAGGCAG GCACAACTAA AGCCGGGTAC CCTTATAATT   6827
```

101

```
ATAACACCAC TGCTAGCGAC CAACTGCTTG TCGAGAATGC CGCCGGGCAC CGGGTCGCTA    6887

TTTCCACTTA CACCACTAGC CTGGGTGCTG GTCCCGTCTC CATTTCTGCG GTTGCCGTTT    6947

TAGCCCCCCA CTCTGCGCTA GCATTGCTTG AGGATACCTT GGACTACCCT GCCCGCGCCC    7007

ATACTTTTGA TGATTTCTGC CCAGAGTGCC GCCCCCTTGG CCTTCAGGGC TGCGCTTTCC    7067

AGTCTACTGT CGCTGAGCTT CAGCGCCTTA AGATGAAGGT GGGTAAAACT CGGGAGTTGT    7127

AG TTTATTTGCT TGTGCCCCCC TTCTTTCTGT TGCTTATTTC TCATTTCTGC            7179

GTTCCGCGCT CCCTGA                                                    7195
```

a fourth sequence (SEQ ID NO.10):

```
GCCATGGAGG CCCACCAGTT CATTAAGGCT CCTGGCATCA CTACTGCTAT TGAGCAAGCA      60

GCTCTAGCAG CGGCCAACTC CGCCCTTGCG AATGCTGTGG TGGTCCGGCC TTTCCTTTCC     120

CATCAGCAGG TTGAGATCCT TATAAATCTC ATGCAACCTC GGCAGCTGGT GTTTCGTCCT     180

GAGGTTTTTT GGAATCACCC GATTCAACGT GTTATACATA ATGAGCTTGA GCAGTATTGC     240

CGTGCTCGCT CGGGTCGCTG CCTTGAGATT GGAGCCCACC CACGCTCCAT TAATGATAAT     300

CCTAATGTCC TCCATCGCTG CTTTCTCCAC CCCGTCGGCC GGGATGTTCA GCGCTGGTAC     360

ACAGCCCCGA CTAGGGGACC TGCGGCGAAC TGTCGCCGCT CGGCACTTCG TGGTCTGCCA     420

CCAGCCGACC GCACTTACTG TTTTGATGGC TTTGCCGGCT GCGGTTTTGC CGCCGAGACT     480

GGTGTGGCTC TCTATTCTCT CCATGACTTG CAGCCGGCTG ATGTTGCCGA GGCGATGGCT     540

CGCCACGGCA TGACCCGCCT TTATGCAGCT TTCCACTTGC CTCCAGAGGT GCTCCTGCCT     600

CCTGGCACCT ACCGGACATC ATCCTACTTG CTGATCCAGG ATGGTAAGCG CGCGGTTGTC     660

ACTTATGAGG GTGACACTAG CGCCGGTTAC AATCATGATG TTGCCACCCT CCGCACATGG     720

ATCAGGACAA CTAAGGTTGT GGGTGAACAC CCTTTGGTGA TCGAGCGGGT GCGGGGTATT     780

GGCTGTCACT TTGTGTTGTT GATCACTGCG GCCCCTGAGC CCTCCCCGAT GCCCTACGTT     840

CCTTACCCGC GTTCGACGGA GGTCTATGTC CGGTCTATCT TTGGGCCCGG CGGGTCCCCG     900

TCGCTGTTCC CGACCGCTTG TGCTGTCAAG TCCACTTTTC ACGCCGTCCC CACGCACATC     960

TGGGACCGTC TCATGCTCTT TGGGGCCACC CTCGACGACC AGGCCTTTTG CTGCTCCAGG    1020

CTTATGACGT ACCTTCGTGG CATTAGCTAT AAGGTAACTG TGGGTGCCCT GGTCGCTAAT    1080

GAAGGCTGGA ATGCCACCGA GGATGCGCTC ACTGCAGTTA TTACGGCGGC TTACCTCACA    1140

ATATGTCATC AGCGTTATTT GCGGACCCAG GCGATTTCTA AGGGCATGCG CCGGCTTGAG    1200
```

CTTGAACATG CTCAGAAATT TATTTCACGC CTCTACAGCT GGCTATTTGA GAAGTCAGGT  1260

CGTGATTACA TCCCAGGCCG CCAGCTGCAG TTCTACGCTC AGTGCCGCCG CTGGTTATCT  1320

GCCGGGTTCC ATCTCGACCC CCGCACCTTA GTTTTTGATG AGTCAGTGCC TTGTAGCTGC  1380

CGAACCACCA TCCGGCGGAT CGCTGGAAAA TTTTGCTGTT TTATGAAGTG GCTCGGTCAG  1440

GAGTGTTCTT GTTTCCTCGA GCCCGCCGAG GGGCTGGCGG GCGACCAAGG TCATGACAAT  1500

GAGGCCTATG AAGGCTCTGA TGTTGATACT GCTGAGCCTG CCACCCTAGA CATTACAGGC  1560

TCATACATCG TGGATGGTCG GTCTCTGCAA ACTGTCTATC AAGCTCTCGA CCTGCCAGCT  1620

GACCTGGTAG CTCGCGCAGC CCGACTGTCT GCTACAGTTA CTGTTACTGA AACCTCTGGC  1680

CGTCTGGATT GCCAAACAAT GATGGGCAAT AAGACTTTTC TCACTACCTT TGTTGATGGG  1740

GCACGCCTTG AGGTTAACGG GCCTGAGCAG CTTAACCTCT CTTTTGACAG CCAGCAGTGT  1800

AGTATGGCAG CCGGCCCGTT TTGCCTCACC TATGCTGCCG TAGATGGCGG GCTGGAAGTT  1860

CATTTTTCCA CCGCTGGCCT CGAGAGCCGT GTTGTTTTCC CCCCTGGTAA TGCCCCGACT  1920

GCCCCGCCGA GTGAGGTCAC CGCCTTCTGC TCAGCTCTTT ATAGGCACAA CCGGCAGAGC  1980

CAGCGCCAGT CGGTTATTGG TAGTTTGTGG CTGCACCCTG AAGGTTTGCT CGGCCTGTTC  2040

CCGCCCTTTT CACCCGGGCA TGAGTGGCGG TCTGCTAACC CATTTTGCGG CGAGAGCACG  2100

CTCTACACCC GCACTTGGTC CACAATTACA GACACACCCT TAACTGTCGG GCTAATTTCC  2160

GGTCATTGG ATGCTGCTCC CCACTCGGGG GGGCCACCTG CTACTGCCAC AGGCCCTGCT  2220

GTAGGCTCGT CTGACTCTCC AGACCCTGAC CCGCTACCTG ATGTTACAGA TGGCTCACGC  2280

CCCTCTGGGG CCCGTCCGGC TGGCCCCAAC CCGAATGGCG TTCCGCAGCG CCGCTTACTA  2340

CACACCTACC CTGACGGCGC TAAGATCTAT GTCGGCTCCA TTTTCGAGTC TGAGTGCACC  2400

TGGCTTGTCA ACGCATCTAA CGCCGGCCAC CGCCCTGGTG GCGGGCTTTG TCATGCTTTT  2460

TTTCAGCGTT ACCCTGATTC GTTTGACGCC ACCAAGTTTG TGATGCGTGA TGGTCTTGCC  2520

GCGTATACCC TTACACCCCG GCCGATCATT CATGCGGTGG CCCCGGACTA TCGATTGGAA  2580

CATAACCCCA AGAGGCTCGA GGCTGCCTAC CGCGAGACTT GCGCCCGCCG AGGCACTGCT  2640

GCCTATCCAC TCTTAGGCGC TGGCATTTAC CAGGTGCCTG TTAGTTTGAG TTTTGATGCC  2700

TGGGAGCGGA ACCACCGCCC GTTTGACGAG CTTTACCTAA CAGAGCTGGC GGCTCGGTGG  2760

TTTGAATCCA ACCGCCCCGG TCAGCCCACG TTGAACATAA CTGAGGATAC CGCCCGTGCG  2820

GCCAACCTGG CCCTGGAGCT TGACTCCGGG AGTGAAGTAG GCCGCGCATG TGCCGGGTGT  2880

103

```
AAAGTCGAGC CTGGCGTTGT GCGGTATCAG TTTACAGCCG GTGTCCCCGG CTCTGGCAAG   2940

TCAAAGTCCG TGCAACAGGC GGATGTGGAT GTTGTTGTTG TGCCCACTCG CGAGCTTCGG   3000

AACGCTTGGC GGCGCCGGGG CTTTGCGGCA TTCACTCCGC ACACTGCGGC CCGTGTCACT   3060

AGCGGCCGTA GGGTTGTCAT TGATGAGGCC CCTTCGCTCC CCCCACACTT GCTGCTTTTA   3120

CATATGCAGC GTGCTGCATC TGTGCACCTC CTTGGGGACC CGAATCAGAT CCCCGCCATA   3180

GATTTTGAGC ACACCGGTCT GATTCCAGCA ATACGGCCGG AGTTGGTCCC GACTTCATGG   3240

TGGCATGTCA CCCACCGTTG CCCTGCAGAT GTCTGTGAGT TAGTCCGTGG TGCTTACCCT   3300

AAAATCCAGA CTACAAGTAA GGTGCTCCGT TCCCTTTTCT GGGGAGAGCC AGCTGTCGGC   3360

CAGAAGCTAG TGTTCACACA GGCTGCTAAG GCCGCGCACC CCGGATCTAT AACGGTCCAT   3420

GAGGCCCAGG GTGCCACTTT TACCACTACA ACTATAATTG CAACTGCAGA TGCCCGTGGC   3480

CTCATACAGT CCTCCCGGGC TCACGCTATA GTTGCTCTCA CTAGGCATAC TGAAAAATGT   3540

GTTATACTTG ACTCTCCCGG CCTGTTGCGT GAGGTGGGTA TCTCAGATGC CATTGTTAAT   3600

AATTTCTTCC TTTCGGGTGG CGAGGTTGGT CACCAGAGAC CATCGGTCAT TCCGCGAGGC   3660

AACCCTGACC GCAATGTTGA CGTGCTTGCG GCGTTTCCAC CTTCATGCCA AATAAGCGCC   3720

TTCCATCAGC TTGCTGAGGA GCTGGGCCAC CGGCCGGCGC CGGTGGCGGC TGTGCTACCT   3780

CCCTGCCCTG AGCTTGAGCA GGGCCTTCTC TATCTGCCAC AGGAGCTAGC CTCCTGTGAC   3840

AGTGTTGTGA CATTTGAGCT AACTGACATT GTGCACTGCC GCATGGCGGC CCCTAGCCAA   3900

AGGAAAGCTG TTTTGTCCAC GCTGGTAGGC CGGTATGGCA GACGCACAAG GCTTTATGAT   3960

GCGGGTCACA CCGATGTCCG CGCCTCCCTT GCGCGCTTTA TTCCCACTCT CGGGCGGGTT   4020

ACTGCCACCA CCTGTGAACT CTTTGAGCTT GTAGAGGCGA TGGTGGAGAA GGGCCAAGAC   4080

GGTTCAGCCG TCCTCGAGTT GGATTTGTGC AGCCGAGATG TCTCCCGCAT AACCTTTTTC   4140

CAGAAGGATT GTAACAAGTT CACGACCGGC GAGACAATTG CGCATGGCAA AGTCGGTCAG   4200

GGTATCTTCC GCTGGAGTAA GACGTTTTGT GCCCTGTTTG GCCCGTGGTT CCGTGCGATT   4260

GAGAAGGCTA TTCTATCCCT TTTACCACAA GCTGTGTTCT ACGGGGATGC TTATGACGAC   4320

TCAGTATTCT CTGCTGCCGT GGCTGGCGCC AGCCATGCCA TGGTGTTTGA AAATGATTTT   4380

TCTGAGTTTG ACTCGACTCA GAATAACTTT TCCCTAGGTC TTGAGTGCGC CATTATGGAA   4440

GAGTGTGGTA TGCCCCAGTG GCTTGTCAGG TTGTACCATG CCGTCCGGTC GGCGTGGATC   4500

CTGCAGGCCC CAAAAGAGTC TTTGAGAGGG TTCTGGAAGA AGCATTCTGG TGAGCCGGGC   4560
```

104

```
AGCTTGCTCT GGAATACGGT GTGGAACATG GCAATCATTG CCCATTGCTA TGAGTTCCGG      4620

GACCTCCAGG TTGCCGCCTT CAAGGGCGAC GACTCGGTCG TCCTCTGTAG TGAATACCGC      4680

CAGAGCCCAG GCGCCGGTTC GCTTATAGCA GGCTGTGGTT TGAAGTTGAA GGCTGACTTC      4740

CGGCCGATTG GGCTGTATGC CGGGGTTGTC GTCGCCCCGG GGCTCGGGGC CCTACCCGAT      4800

GTCGTTCGAT TCGCCGGACG GCTTTCGGAG AAGAACTGGG GGCCTGATCC GGAGCGGGCA      4860

GAGCAGCTCC GCCTCGCCGT GCAGGATTTC CTCCGTAGGT TAACGAATGT GGCCCAGATT      4920

TGTGTTGAGG TGGTGTCTAG AGTTTACGGG GTTTCCCCGG GTCTGGTTCA TAACCTGATA      4980

GGCATGCTCC AGACTATTGG TGATGGTAAG GCGCATTTTA CAGAGTCTGT TAAGCCTATA      5040

CTTGACCTTA CACACTGAAT TATGCACCGG TCTGAATGAA TAACATGTGG TTTGCTGCGC      5100

CCATGGGTTC GCCACCATGC GCCCTAGGCC TCTTTTGCTG TTGTTCCTCT TGTTTCTGCC      5160

TATGTTGCCC GCGCCACCGA CCGGTCAGCC GTCTGGCCGC CGTCGTGGGC GGCGCAGCGG      5220

CGGTACCGGC GGTGGTTTCT GGGGTGACCG GGTTGATTCT CAGCCCTTCG CAATCCCCTA      5280

TATTCATCCA ACCAACCCCT TTGCCCCAGA CGTTGCCGCT GCGTCCGGGT CTGGACCTCG      5340

CCTTCGCCAA CCAGCCCGGC CACTTGGCTC CACTTGGCGA GATCAGGCCC AGCGCCCCTC      5400

CGCTGCCTCC CGTCGCCGAC CTGCCACAGC CGGGGCTGCG GCGCTGACGG CTGTGGCGCC      5460

TGCCCATGAC ACCTCACCCG TCCCGGACGT TGATTCTCGC GGTGCAATTC TACGCCGCCA      5520

GTATAATTTG TCTACTTCAC CCGTGACATC CTCTGTGGCC TCTGGCACTA ATTTAGTCCT      5580

GTATGCAGCC CCCCTTAATC CGCCTCTGCC GCTGCAGGAC GGTACTAATA CTCACATTAT      5640

GGCCACAGAG GCCTCCAATT ATGCACAGTA CCGGGTTGCC CGCGCTACTA TCCGTTACCG      5700

GCCCCTAGTG CCTAATGCAG TTGGAGGCTA TGCTATATCC ATTTCTTTCT GGCCTCAAAC      5760

AACCACAACC CCTACATCTG TTGACATGAA TTCCATTACT TCCACTGATG TCAGGATTCT      5820

TGTTCAACCT GGCATAGCAT CTGAATTGGT CATCCCAAGC GAGCGCCTTC ACTACCGCAA      5880

TCAAGGTTGG CGCTCGGTTG AGACATCTGG TGTTGCTGAG GAGGAAGCCA CCTCCGGTCT      5940

TGTCATGTTA TGCATACATG GCTCTCCAGT TAACTCCTAT ACCAATACCC CTTATACCGG      6000

TGCCCTTGGC TTACTGGACT TTGCCTTAGA GCTTGAGTTT CGCAATCTCA CCACCTGTAA      6060

CACCAATACA CGTGTGTCCC GTTACTCCAG CACTGCTCGT CACTCCGCCC GAGGGGCCGA      6120

CGGGACTGCG GAGCTGACCA CAACTGCAGC CACCAGGTTC ATGAAAGATC TCCACTTTAC      6180

CGGCCTTAAT GGGGTAGGTG AAGTCGGCCG CGGGATAGCT CTAACATTAC TTAACCTTGC      6240
```

105

TGACACGCTC CTCGGCGGGC TCCCGACAGA ATTAATTTCG TCGGCTGGCG GGCAACTGTT  6300

TTATTCCCGC CGGGTTGTCT CAGCCAATGG CGAGCCAACC GTGAAGCTCT ATACATCAGT  6360

GGAGAATGCT CAGCAGGATA AGGGTGTTGC TATCCCCCAC GATATCGATC TTGGTGATTC  6420

GCGTGTGGTC ATTCAGGATT ATGACAACCA GCATGAGCAG GATCGGCCCA CCCCGTCGCC  6480

TGCGCCATCT CGGCCTTTTT CTGTTCTCCG AGCAAATGAT GTACTTTGGC TGTCCCTCAC  6540

TGCAGCCGAG TATGACCAGT CCACTTACGG GTCGTCAACT GGCCCGGTTT ATATCTCGGA  6600

CAGCGTGACT TTGGTGAATG TTGCGACTGG CGCGCAGGCC GTAGCCCGAT CGCTTGACTG  6660

GTCCAAAGTC ACCCTCGACG GGCGGCCCCT CCCGACTGTT GAGCAATATT CCAAGACATT  6720

CTTTGTGCTC CCCCTTCGTG GCAAGCTCTC CTTTTGGGAG GCCGGCACAA CAAAAGCAGG  6780

TTATCCTTAT AATTATAATA CTACTGCTAG TGACCAGATT CTGATTGAAA ATGCTGCCGG  6840

CCATCGGGTC GCCATTTCAA CCTATACCAC CAGGCTTGGG GCCGGTCCGG TCGCCATTTC  6900

TGCGGCCGCG GTTTTGGCTC CACGCTCCGC CCTGGCTCTG CTGGAGGATA CTTTTGATTA  6960

TCCGGGGCGG GCGCACACAT TTGATGACTT CTGCCCTGAA TGCCGCGCTT TAGGCCTCCA  7020

GGGTTGTGCT TTCCAGTCAA CTGTCGCTGA GCTCCAGCGC CTTAAAGTTA AGGTGGGTAA  7080

AACTCGGGAG TTGTAGTTTA TTTGGCTGTG CCCACCTACT TATATCTGCT GATTTCCTTT  7140

ATTTCCTTTT TCTCGGTCCC GCGCTCCCTG A  7171

or a fifth sequence (SEQ ID NO.12):
CGGGCCCCGT ACAGGTCACA ACCTGTGAGT TGTACGAGCT AGTGGAGGCC ATGGTCGAGA   60

AAGGCCAGGA TGGCTCCGCC GTCCTTGAGC TCGATCTCTG CAACCGTGAC GTGTCCAGGA  120

TCACCTTTTT CCAGAAAGAT TGCAATAAGT TCACCACGGG AGAGACCATC GCCCATGGTA  180

AAGTGGGCCA GGGCATTTCG GCCTGGAGTA AGACCTTCTG TGCCCTTTTC GGCCCCTGGT  240

TCCGTGCTAT TGAGAAGGCT ATTCTGGCCC TGCTCCCTCA GGGTGTGTTT TATGGGGATG  300

CCTTTGATGA CACCGTCTTC TCGGCGCGTG TGGCCGCAGC AAAGGCGTCC ATGGTGTTTG  360

AGAATGACTT TTCTGAGTTT GACTCCACCC AGAATAATTT TTCCCTGGGC CTAGAGTGTG  420

CTATTATGGA GAAGTGTGGG ATGCCGAAGT GGCTCATCCG CTTGTACCAC CTTATAAGGT  480

CTGCGTGGAT CCTGCAGGCC CCGAAGGAGT CCCTGCGAGG GTGTTGGAAG AAACACTCCG  540

GTGAGCCCGG CACTCTTCTA TGGAATACTG TCTGGAACAT GGCCGTTATC ACCCATTGTT  600

ACGATTTCCG CGATTTGCAG GTGGCTGCCT TTAAAGGTGA TGATTCGATA GTGCTTTGCA  660

GTGAGTACCG TCAGAGTCCA GGGGCTGCTG TCCTGATTGC TGGCTGTGGC TTAAAGCTGA  720

```
AGGTGGGTTT CCGTCCGATT GGTTTGTATG CAGGTGTTGT GGTGACCCCC GGCCTTGGCG      780

CGCTTCCCGA CGTCGTGCGC TTGTCCGGCC GGCTTACTGA GAAGAATTGG GGCCCTGGCC      840

CTGAGCGGGC GGAGCAGCTC CGCCTTGCTG TGCG                                  874
```

or a sequence complementary thereto.


　　　14.  A kit comprising, in a container or separate
containers, a pair of single-strand primers derived
from nonhomologous regions of opposite strands of a
DNA duplex fragment derived from an enterically
transmitted viral hepatitis agent whose genome
contains a region which is homologous to the 1.33 kb
DNA EcoRI insert present in plasmid pTZKF1(ET1.1)
carried in E. coli strain BB4 and having ATCC deposit
no. 67717.


　　　15.  The kit of claim 15, which are derived from
opposite strands of the EcoRI duplex insert in said
plasmid.


　　　16.  A method for detecting the presence of an
enterically transmitted nonA/nonB hepatitis viral
agent in a biological sample, comprising
　　　preparing a mixture of duplex DNA fragments
derived from the sample,
　　　denaturing the duplex fragments,
　　　adding to the denatured DNA fragments, a pair of
single-strand primers derived from nonhomologous
regions of opposite strands of a DNA duplex fragment
derived from an enterically transmitted viral
hepatitis agent whose genome contains a region which
is homologous to the 1.33 kb DNA EcoRI insert present
in plasmid pTZKF1(ET1.1) carried in E. coli strain
BB4, and having ATCC deposit no. 67717,
　　　hybridizing said primers to homologous-sequence
region of opposite strands of such duplex DNA

107

fragments derived from enterically transmitted nonA/nonB hepatitis agent,

reacting the primed fragment strands with DNA polymerase in the presence of DNA nucleotides, to form new DNA duplexes containing the primer sequences, and

repeating said denaturing, adding, hybridizing and reacting steps, until a desired degree of amplification of sequences is achieved.

17. The method of claim 16, wherein the primers are derived from opposite strands of the EcoRI duplex insert in said plasmid.

18. The method of claim 16, for detecting the presence of viral agent in a sample of cultured cells infected with the agent.

19. A vaccine for immunizing an individual against enterically transmitted nonA/nonB hepatitis viral agent comprising, in a pharmacologically acceptable adjuvant, a recombinant protein derived from an enterically transmitted nonA/nonB viral hepatitis agent whose genome contains a region which is homologous to the 1.33 kb DNA EcoRI insert present in plasmid pTZ-RF1(ET1.1) carried in *E. coli* strain BB4, and having ATCC deposit no. 67717.

20. The vaccine of claim 19, wherein the protein is derived from the EcoRI insert in said plasmid.

21. A vaccine for immunizing an individual against HEV comprising, in a pharmacologically acceptable adjuvant, a protein encoded by genetic sequence 406.3-2 or 406.4-2 or a fragment thereof.

22. In a method of isolating an enterically transmitted nonA/nonB viral agent or a nucleic acid fragment produced by the agent, an improvement which

108

comprises: utilizing, as a source of said agent, bile obtained from a human or cynomolgus monkey having an active infection of enterically transmitted non-A/non-B hepatitis.

5

23.  The method of claim 22, wherein the bile is obtained from an infected cynomolgus monkey.

24.  Human polyclonal anti-serum obtained from a
10  human immunized with a protein derived form an enterically transmitted non-A/non-B viral hepatitis agent whose genome contains a region which is homologous to the 1.33 kb DNA EcoRI insert present in plasmid pTZKF1(ET1.1) carried in E. coli strain BB4
15  and having ATCC deposit no. 67717.

ADD
A17

109